BRIDGING THE TEST GAP:
FINDINGS FROM A FIVE STATE GAP ANALYSIS
FOR THE AMERICAN DIPLOMA PROJECT

THE VIEW ACROSS THE STATES

SUMMARY OF FINDINGS

General findings:

- Some states have narrower gaps than others, but none is perfectly aligned;
- Gaps are clearest in mathematics, especially Algebra 2 and Statistics;
- Gaps are smallest in reading; several current tests can provide useful information to higher education; and
- Writing assessments for high school are sometimes more sophisticated than writing assessments for college placement.

Reading comprehension

- The reading passages on the SAT 1, ACT, and high school tests in three ADP states (Massachusetts, Indiana and Texas) were appropriate for college readiness;
- High school tests in Indiana, Kentucky, Massachusetts and Texas made more frequent use of open-response items for reading, which added more complexity to the tests;
- High school tests, particularly Nevada and Kentucky, tended to cover a wider range of genres, including functional or document reading; and
- Some high school exams and placement tests featured texts that were either too short or too low level to determine college readiness.

Writing

- Every state had a high school on-demand writing test. These could either provide useful information for higher education now, or could be easily strengthened so they will;
- High school writing assessments were often more sophisticated than placement tests;
- There is no writing on ACT or SAT 1.
- Criteria for evaluating writing for placement are typically ambiguous and vary from institution to institution.

Mathematics

- All tests emphasized algebra & functions, but college admissions and placement tests addressed this mathematics strand almost exclusively;
- College placement tests and the ACT assess Algebra 2, which is lacking on most state high school assessments and the SAT 1;
- Most high school tests assess statistics and geometry, which are lacking on the college tests.
- College admissions and placement tests are predominately procedural; high school tests tend to also include items testing students' conceptual understanding and problem solving skill.

BRIDGING THE TEST GAP*:*
FINDINGS FROM A FIVE STATE GAP ANALYSIS
FOR THE AMERICAN DIPLOMA PROJECT

THE VIEW ACROSS THE STATES

Prepared by the Education Trust, Inc., July 2002

Introduction

Five states—Indiana, Kentucky, Massachusetts, Nevada and Texas—were selected to be partners in the American Diploma Project (ADP) along with four national organizations: Achieve, Inc., the Education Trust, the Fordham Foundation and the National Alliance of Business. ADP is an initiative through which partner states will begin to align their high school exit standards with the knowledge and skills needed for success in state postsecondary institutions and employment in high-performance jobs.

As partners in the American Diploma Project, the five ADP states made two commitments:

- to use their present high school assessment data in the processes for college admissions/placement and employment; and
- to develop plans for aligning the standards and assessments between the two education sectors—K-12 and higher education—in their state so that students can be confident that their diploma prepares them for a range of postsecondary options.

The first step in deciding where benchmarks should be set is determining where they currently are. K-12 standards are a matter of public record. But the standards embedded in college admissions tests and especially those used by universities for placement into credit-bearing work aren't as well known by teachers, parents and students, or for that matter by college faculty. Even less is known about how these standards and tests relate to each other. Conversations across education sectors are uncommon. In most states, the standards and tests used by one sector were designed with little attention to the needs of the other. Consequently, there is typically little or no coherence across systems. At best, students are left confused trying to determine what knowledge and skills are important to their futures; at worst, they find themselves woefully ill-prepared for college-level work.

Through the auspices of ADP, faculty teams from K-12 and higher education representing each state met earlier this year to examine the tests taken by a hypothetical student who is leaving high school and entering postsecondary education in their state. The reviewers analyzed the tests with this overarching question in mind:

Does this test reveal some useful information to higher education?

The teams used a common analytical framework developed for the Diploma Project to describe the content and demands of the tests relevant to their particular states. They then identified gaps between them as well as the knowledge that both sectors value. The analysis was limited to mathematics and English language arts.

The faculty teams had primary responsibility for conducting the analysis for their respective states. Their findings are presented in individual state reports prepared by Education Trust staff who shared drafts with state coordinators for the Diploma Project in order to make sure that the reports accurately represent the faculty teams' analyses. The Ed Trust has synthesized the teams' findings and we offer our own analysis in this cross-state view to help states as they think through the issues related to the project. Both this and the state reports will provide a foundation for larger conversations in the ADP states about how to bridge any gaps and send a coherent message to students and schools.

A word about the assessments

The five partner states of the American Diploma Project provide different images for designing and using end-of-high school tests. Massachusetts, Texas and Nevada have each developed comprehensive graduation exams that assess students' mastery of the core content their states require for a diploma. These tests address a range of concepts and skills that students are expected to accumulate through their secondary school years. Kentucky also administers comprehensive tests to all high school students with one distinction: the Kentucky assessments are primarily used to measure school performance and are not required as a condition for graduation. Although Indiana currently has an exit exam, the state is developing a series of end-of-course (EOC) exams that state leaders anticipate will eventually be used as part of their graduation requirements. Unlike the comprehensive tests in the other four states, Indiana's EOCs are focused on the content learned in specific high school courses. Students will likely be required to pass a combination of EOCs for a diploma, possibly with some limited choices about which tests they can take to graduate.

The ADP states also differ according to when they have students take the tests. The so-called exit exams are actually given to students long before they exit high school at the end of their senior year. Massachusetts students initially take them in grade 10. Nevada students first take them in grade 11. Texas high schoolers now see the exit exam first in grade 10, but their peers will take the revised test initially in grade 11. EOCs are by definition given at the grade level the course is taken, so in Indiana, students will take a battery of state assessments beginning early in high school and continuing each year. Kentucky's comprehensive tests are administered on a staggered schedule: students take the ELA test in grade 10, mathematics in grade 11, and submit a writing portfolio in grade 12.

The practice of giving a high school graduation test as early as grade 10 may seem anachronistic, but it's an indication of how muddy the line is between secondary and postsecondary. Surely, one reason for testing earlier in high school rather than at the end is to extend ample time and opportunity to students to pass an exam that will have definite consequences for their futures. Another reason is purely practical. States can take several months to process student scores and get them back to schools. If the test is given too close to the end of the senior year, diplomas and commencement ceremonies could be kept hanging in the balance while administrators await results.

Even if diplomas aren't at stake, it can still make sense to administer the tests early. Test results provide valuable feedback to students, teachers and counselors about their academic progress. The earlier they get this information, the easier it is for students to get help where they need it, and for schools to continually improve their academic programs.

Regardless of the reasons for early testing policies, they leave states with a conundrum: if state accountability is targeted at grades 10 or 11, what is the value gained by the senior year?

Higher education testing requirements also blur the definition of what "end of high school" means. Colleges and universities in all the ADP states admit students based in part on their scores on national college admissions tests, either the SAT 1 or ACT. Aspiring applicants typically take these tests in grade 11, and increasing numbers are taking them in grade 10.

In truth, college placement tests were the only exams in this analysis that are typically taken at the point when high school ends and college begins. But even here there are exceptions. Kentucky, for example, offers high school sophomores and juniors the chance to take a math placement test for diagnostic purposes. Students then use the results to plan their schedules so they will be better prepared for success when they get to college.

All of which establishes a context for viewing the assessments in this study. Although the tests in this study are all theoretically targeted to students at the same transitional time in their lives—preparing for life after high school—they are designed for different purposes and taken at different points in high school. Some content differences among tests should therefore be expected. At the same time, there should be enough coherence between them to provide some useful information to higher education about a student's readiness for college-level work, even if there is not a precise match.

Interpreting cross-state findings: A caution

Trying to characterize how "difficult" a test is is not a judgment based on any single dimension. Test questions can be "hard" because the level of content is high or because the test-taker has to go through multiple steps in order to come up with an answer. "Easy" questions can address low-level content or the answers can be very straightforward and obvious.

But the relationship is not always clear or predictable. For example, test questions can be very straightforward about high-level content. *Hamlet* is typically taught in the senior year of high school, so it would be an arguably "high-level" topic. But suppose this was the test question:

> *Hamlet* is best described as which of the following terms?
> A. tragedy
> B. comedy
> C. musical
> D. talk show

This item demands no mental effort beyond simple fact retrieval. Is this a hard or easy question? Of course it would be hard if you were never taught anything about Shakespeare. But students who studied Shakespeare in their high school curriculum could answer this question almost without thinking, even if they had not read the play.

On the other hand, tests can ask very complex questions about low-level content. Consider this grade 12 NAEP item:

> A rectangular pool 24 feet long, 8 feet wide, and 4 feet deep is filled with water. Water is leaking from the pool at the rate of 0.4 cubic foot per minute. At this rate, how many hours will it take for the water level to drop 1 foot?

A. 4
B. 8
C. 12
D. 16
E. 32

This is fundamentally an arithmetic problem, using math that students probably learned by grade 7. But it also has several moving parts that the test-taker has to put together in order to come up with the correct response. In addition, students need a conceptual understanding beyond just knowing the algorithms. And they need to be skillful at problem solving. Despite the basic computation embedded in the problem, this item proved to be very difficult for high school seniors. Only one in four got it right.[1]

Reviewers took the relationship among all the test components into account when judging its usefulness for higher education. Readers of this report should use the same care. Just because a particular test includes a few middle school-level topics or some simple items, it should not automatically be assumed that it is a low-level assessment. Sometimes it's simply worthwhile to know whether a student leaving high school can still apply grammar rules learned in grade 8, for example. If, however, such items predominate, the test is probably too low for either college readiness or high school graduation. A good rule of thumb is to look for a reasonable balance within and between categories, and make sure that the high end is high enough.

INTERPRETING FINDINGS: A good rule of thumb is to look for a reasonable balance within and between categories, and make sure that the high end is high enough.

General findings

As mentioned, the state reviewers attempted to find out if a particular test, regardless of its original purpose, could provide information about a student's knowledge and skills that higher education would find useful for admissions or placement decisions. "Usefulness" was interpreted broadly and could mean anything from accepting high scores on the high school exit exam for college placement purposes to merely allowing students to voluntarily include their scores on their college application.

The college tests were as open to the question of usefulness as the high school tests. The answers were not always predictable. Many high school tests addressed content and skills at a level that could be meaningful to admissions officers. On the other hand, reviewers found some college admissions and placement tests to be too narrow to show student's readiness for some aspects of college-level work.

The good news is that in every ADP state, reviewers found something valuable with which to build a foundation for secondary-postsecondary alignment. Even so, not one state in the project had a testing system that is well aligned across sectors in every subject. But some states came close, particularly in the area of English language arts.

In general, the college tests tended to feature content at a higher grade level than the high school assessments, but the high school exams tended to cover a broader range of content than the college tests did. The tests also differed by how demanding the items were, that is, the extent to which items require higher levels of thinking or manipulation to produce a correct response. Some high school tests had a higher proportion of demanding items than

[1] US Department of Education, NAEP Test Questions, www.nces.ed.gov/nationsreportcard.

the college tests, even though the concepts themselves may be a lower grade level. For example, the high school mathematics assessments were more likely to address students' conceptual understanding compared to the college entry and placement tests reviewed. Likewise, high school reading tests generally had more "complex" items, which require test-takers to go through multiple steps to arrive at the answer. Clearly, the trend to include open-response items on high school tests of reading and mathematics (all ADP state assessments except Nevada's had such items) is adding to the level of demand for students.

Looking across the states, we can make some general observations:

- ***Reading.*** While there were some gaps between the high school and college tests of reading comprehension, the two sectors are closer in this subject than in either mathematics or writing. Reviewers for Indiana, Massachusetts and Texas found that the high school assessments in their states were particularly close to being aligned to the skills needed for college.

- ***Writing.*** Reviewers in every state thought that their current high school writing assessment could either be considered useful to higher education now, depending on the scoring criteria, or be strengthened fairly easily so that it could be. In many cases, it was because the college assessments were found to provide little information about students' readiness for college-level writing. The SAT 1 and ACT are the most egregious examples since neither exam even asks for students to write. In contrast, all of the ADP states have an on-demand writing task as part of their high school assessments.

- ***Mathematics.*** The gap between high school exit and college placement was clearest and widest in mathematics. In addition, there were gaps on both sides. The high school assessments in every state except Indiana and the SAT I had little or no Algebra 2 content. On the other hand, college placement tests were less likely to feature Geometry than either the admissions tests or the high school assessments. Statistics items were nearly or completely absent from the college tests and from the high school assessments in Indiana and Texas.[2]

We offer a more detailed cross-state view of the tests by subject area below.

Reading

Across assessments, the tests of reading comprehension followed the same basic pattern: a short reading passage or sometimes two (anywhere from 100-800 words) followed by a series of questions relating to the reading. The questions were either all multiple choice or a combination of multiple-choice and short open-response.

The grade level and complexity of the reading passages have a strong relationship to evaluating the difficulty of the test. This is because the nature of reading skills is generally unchanged from the time a student has made the leap from learning how to read to using reading to learn. Students from grade 3 on make inferences, interpret and analyze what they

[2] The College Board has recently announced plans to overhaul the SAT 1. Among the Board's recommendations are to require a written essay and to include more Algebra 2.

read. The skills typically grow in response to the sophistication of the texts they interact with. Therefore, test items that ask students to identify the main idea of a reading passage, for example, can be either high- or low-level depending on the vocabulary, complexity, nuances, etc. of the text.

For this reason, when judging the relative merits of the test, reviewers examined the interaction of reading passages and questions. For example, a rich middle school level text that is followed with a series of items demanding inference, interpretation and analysis could be a good assessment of critical reading skills at an upper high school level. Similarly, simple recognition questions could reveal whether students understand the literal meaning of complicated, unfamiliar texts like those they will surely encounter in college. Again, readers are reminded that the inclusion of *some* middle school level or simple, recall items can be appropriate, as long as there is sufficient interaction with higher level and/or more complex content on the test as a whole.

The table on page 6 shows the distribution of items by category for each English language arts test.

The major gaps in reading are:

- Reviewers found that the texts featured on the SAT 1-verbal, the ACT-reading, and the high school assessments in Indiana (grade 12 EOC), Massachusetts and Texas were overall appropriate for college readiness, both in terms of grade level and complexity.

- The state high school reading tests in Indiana, Kentucky, Massachusetts and Texas used open-response questions in addition to multiple-choice. This format offered more opportunities for students to show their critical reading abilities. As a result, the open response items tended to show more complexity.

- The high school tests tended to address a broader range of genres than the college tests. The state assessments for Nevada and Kentucky were particularly noteworthy for including functional or document reading passages in addition to literary and informational texts—genres that are common to all reading tests. Functional/document reading refers to the ability to understand public documents ranging from tax forms and train schedules to legal contracts and technical manuals with charts.

Reviewers found that the skills assessed on computer adaptive placement tests for reading—Compass and Accuplacer—were probably appropriate for college readiness. However, the range of skills was narrow. They also found that the short passages featured on these tests were not at the level of the longer texts characteristic of the college admissions tests and many of the high school assessments.

Gap Analysis - Cross-State View
Reading Comprehension & Editing
(in percent of total test items)

	Grade Level				Cognitive Challenge			Kind of Knowledge				
	4 to 6	6 to 8	9 to 10	11 to 12	Simple	Moderate	Complex	Recall	Proce-dural	Inference	Interpre-tation	Analysis
Indiana												
Core 40 – Gr 11 Reading	0	25	72	3	11	53	36	5	33	15	21	26
Core 40 – Gr 11 Editing	0	67	33	0	75	25	0	0	100	0	0	0
Core 40 – Gr 12	0	26	57	17	40	33	20	11	27	9	23	30
Kentucky												
KCCT- Gr 10	11	71	9	9	43	49	8	28	13	23	23	13
Nevada												
HSPE – Gr 10 Reading	31	43	22	4	59	38	3	49	2	38	4	7
Texas												
TAKS- High School Reading	0	10	52	38	5	57	38	0	0	29	19	52
TAKS - Editing	19	56	25	0	69	31	0	19	62	19	0	0
TASP- Placement Reading	12	43	38	7	24	71	5	26	0	40	10	24
TASP- Editing	20	40	40	0	37.5	62.5	0	40	7.5	52.5	0	0
Massachusetts												
MCAS – Gr 10	1	14	46	39	21	53	26	22	6	31	12	29
MCAS retest	5	38	54	3	28	69	3	29	4	33	19	15
National admissions & placement												
SAT 1	0	12	62	26	23	61	16	21	0	53	14	10
ACT – Reading	11	30	39	20	40	45	15	35	28	33	20	12
ACT – Editing	23	41	25	11	48	40	12	31	39	6	5	19
Compass* (placement)	0	82	12	6	35	53	12	35	0	47	18	0
Accuplacer* (placement)	0	0	52	48	3.5	93	3.5	10	0	86	0	0

***Compass & Accuplacer are computer adaptive tests. The analysis may not be representative of the entire item bank.**

Categories:

Grade level: the point at which the item content is typically taught and learned

Cognitive challenge: the degree of challenge in the item

Kind of knowledge: the cognitive ability that is the primary target of the item

SEE. See "Interpreting cross-state findings", page 4. "Reading & Editing Protocol" for a complete description of categories

Editing

The tests in the analysis varied by how they addressed—or in one case, did not address—editing skills. We defined "editing" tests or sections of English language arts (ELA) tests as a discrete series of questions that ask students to apply rules of grammar, spelling, writing mechanics, etc., as well as general editing. Some of the ELA tests—for example, the Texas Assessment of Student Progress (TASP)—call these sections "writing" exams, even though the format is entirely multiple-choice. For our purposes, we call them "editing" in order to distinguish them from tests that require a written essay.

Editing skills can also be assessed holistically as part of scoring an on-demand writing test. Reviewers examined that approach to testing editing skills as part of the written essay test analysis.

An important thing to keep in mind when considering the editing tests is that middle school is typically the time that students receive the most instruction in grammar and editing rules. Consequently, editing items tend to be classified at the middle-school level. They also tend to target procedural knowledge, or the application of a rule, although reviewers found that one test, the ACT, included a significant proportion of editing items that were inferential and analytical. Most of the editing tests in this study attempted to build some complexity into the items, and push the level of demand beyond simple procedure.

Because the level of editing items tends to be lower than the reading comprehension questions, editing sections of ELA tests, where relevant, were analyzed and reported separately so they would not artificially deflate the grade level recorded for the entire ELA assessment. These assessments are the Texas high school (TAKS) and placement (TASP) exams, the Indiana grade 11 English test, and the ACT. See the "Reading Comprehension & Editing" table on page 8 for a detailed breakdown of editing test items.

Some key observations about editing:

- Three states—Kentucky, Massachusetts and Nevada—assess students' editing skills as part of scoring their on-demand writing test.
- ACT, Accuplacer, Compass, and high school ELA tests in Indiana and Texas featured multiple-choice editing items. As mentioned, because these skills are generally taught in middle school, these items tended to be rated at lower grade levels than reading items.
- SAT 1 had no editing items.

Writing

For this study, "writing" assessments were limited to on-demand writing tasks. The tests are typically one question, or prompt, and students are given a set amount of time to respond in an essay. Test developers often attempt to assess writing independently from curricular content so that scorers won't confuse students' content knowledge with their writing ability. Their challenge is to develop prompts that are interesting and can showcase students' ability to write without privileging students who may have prior knowledge of the topic. For our study, we further wanted to know if the prompt could elicit a response that could show whether or not a student is ready to write at the college level.

Consider this example from a non-ADP state grade 10 assessment:

> Think about a time of year you really like. It could be a season of the year, a special day of the year, or a holiday. Write to your teacher telling why you like this time of year. Be sure to explain why this time is so important to you.

A well-written response to this prompt might reveal students' ability to write grammatical sentences, organize paragraphs, and express thoughts clearly. However, an essay on this topic is not likely to show whether students are able to handle the rigors of academic writing in freshman courses.

Compare the above example to the grade 10 assessment below from another non-ADP state. Articles were provided with the prompt including an overview piece from Knight-Ridder News Service and two opposing opinions from *Insight* magazine.

> You will read a few short articles on the Title IX controversy. Take a position on the issue and write a persuasive editorial for your local newspaper. In your editorial, you must support or oppose the belief that enforcement of Title IX has resulted in a quota system which harms men's athletics. You must include information from each source.

In this example, students must analyze two opposing views, take a position of their own and defend it. They must also cite evidence from the texts provided with the prompt. This prompt clearly demands many of the skills that are crucial to success in college courses.

The prompt is just one aspect of evaluating the demands of a writing test. Even if the prompt is sufficiently rigorous, the scoring criteria can still make the difference between a college-ready essay or a minimal response. Scoring rubrics typically look for students' ability to organize ideas, express them clearly, and provide relevant detail. They also describe how well students use structure and the mechanics of the English language. However, the amount of emphasis on content or on structure is a matter of considerable variation among writing assessments.

Anchor papers remove any remaining ambiguity about judging how good is good enough when scoring essays. Although we sought anchor papers to examine along with the writing assessment, only a few were available for the review. In many cases, notably placement tests, anchor papers did not exist. Therefore, one of the recommendations is for the ADP states to continue the examination of writing assessments used in their state.

The major writing gaps are:

- There is no writing on the SAT 1 or ACT. Despite this fact, many institutions will waive placement tests for freshman-level composition courses based on students' scores on these admissions tests.

- Writing placement tests were typically developed by individual universities. The reviewers found that the many of the prompts wouldn't necessarily elicit a college-ready response. In addition, the criteria for evaluating writing for college placement are typically ambiguous or intuitive, and vary from institution to institution.

  Texas is an exception for administering a common placement test in writing statewide. The Massachusetts higher education system has also established common criteria for

writing across the state, although reviewers thought the criteria were open to interpretation.

- Every ADP state has a high school on-demand writing test. Many of these were more sophisticated than the writing tests used for placement. In every case, reviewers thought that their state's high school writing exam either provides useful information for higher education now, or could be strengthened fairly easily so it will. The determination, however, will be largely dependent on the criteria for scoring.

Mathematics

As mentioned, the content gap between high school and college placement tests was the most pronounced in mathematics, which may explain this subject's reputation as the gatekeeper to college. The gaps were also the easiest to define, particularly in regard to Algebra 2, statistics and geometry.

Algebra and functions comprised the single largest strand of mathematics featured on all the tests in the study, ranging from a low of 27% to 80% of all items. The college placement tests were almost entirely algebra and functions, but Texas and Indiana also put heavy emphasis on this strand in their high school tests.

It's important to note that the placement tests examined in this study were strictly those tests that determine student placement into credit-bearing math courses. There are many more tests available for placing students in the proper course of college-level math. Both Accuplacer and Compass, the computer adaptive placement tests, have forms that can be used for this purpose, and are consequently higher level exams than what is reported here. But this study was not interested in whether a student places into calculus or finite math for English majors; we were only interested in knowing the minimum knowledge and skills needed to stay out of remedial courses.

The major gaps in the mathematics assessments are:

- ***Algebra 2*** Algebra 2 content is a feature of all the college placement tests, yet the high school assessments rarely addressed algebra at that level. The one notable exception was Indiana, which has an EOC for Algebra 2, although the Indiana reviewers thought that the test would have to be more rigorous to be useful for college placement. Kentucky's high school math test is also noteworthy for having 14.6% of its items classified as Algebra 2. Although the Kentucky team thought it was still insufficient to be useful for placement purposes, it's worth noting that it has more Algebra 2 than appears on the SAT 1 (see "Mathematics Grade Level & Demand" table on page 13).

  Because Algebra 2 content represents such a clear gap between tests, we looked at all the Algebra 2 test items across the assessments in the study to describe the concepts and topics that seem to be the most valued by the test designers. In total, the lion's share of Algebra 2 items—57%—involved manipulating, evaluating and simplifying expressions. Polynomials, at 10%, comprised the second most commonly addressed content on the tests. The balance of items were randomly distributed across 12 other topics.

- ***Statistics and geometry***. Both of these mathematical strands are a significant part of several high school exams in the study, but this content is rarely seen on any of the college tests. Statistics made up 15-26% of the items on high school exams in Kentucky, Nevada and Massachusetts. In contrast, it was virtually absent from every placement test, the ACT, and the high school exams in Indiana and Texas. At most these tests had one or two statistics items, and most of them had none at all.

  Likewise, geometry was addressed in significant proportions on the high school tests, especially in Kentucky, Texas and the Indiana EOC in Geometry. But it was lacking on every placement test with the exception of the Texas TASP and Kentucky's early math placement exam (KEMTP) (see "Mathematics Content" table on page 14).

- ***Procedures vs. Conceptual understanding***. There was an observable difference between the two sectors in the kind of knowledge targeted by math items. The vast majority of questions on the college placement tests and on the ACT—between 85 and 100%—were classified as demanding "procedural" knowledge, meaning students must apply an established algorithm or mathematical rule in order to arrive at the answer. The high school tests also emphasized procedural knowledge: at slightly more than half, Kentucky, Massachusetts and Texas had the lowest proportion of such items. But across states, the high school exams had a far greater tendency to assess students' "conceptual" understanding as well. Conceptual items are those that require students to apply their understanding of mathematical concepts to respond correctly. As such, they are *typically* more demanding when addressing the same concept than merely plugging in the algorithm or rule.

  The high school tests in Kentucky, Nevada, Massachusetts and Indiana's EOC in geometry also showed a tendency to include more problem solving items. Problem solving items comprised a small part of these assessments, but in comparison, there was not one problem solving question on any college placement test.

Overall, the high school tests showed a wider range of mathematical content and of the kinds of knowledge students are asked to demonstrate. Nonetheless, the college admissions and placement tests address higher level mathematics, particularly with the emphasis on Algebra 2.

Gap Analysis – Cross-State View
Mathematics Grade Level & Demand
(in percent of total test items)

	Grade Level						Kind of Knowledge			
	Elem	MS	Alg 1	Geom	Alg 2	Other	Recall	Proce-dural	Concep-tual	Problem Solving
Indiana										
Core 40 Alg 1	0	27	73	0	0	0	11	74	16	0
Core 40 Geom	0	35	6	58	0	0	6	74	10	10
Core 40 Alg 2	0	3	36	0	61	0	0	91	6	3
Purdue Placement	0	24	58	0	18	0	3	94	0	3
Kentucky										
KCCT – Gr 11	0	17	44	24.4	14.6	0	12.2	58.5	24.4	4.9
EKU Placement	17.9	14.3	33.9	0	33.9	0	5.4	94.6	0	0
KEMTP Early Plcmnt	0	23	30	13.3	33.3	0	6.7	93.3	0	0
WKU Placement	4	8	52	0	36	0	0	100	0	0
Nevada										
HSPE – math Gr 11	0	44	34	10	2	10	13	65	15	7
Texas										
TAKS – High School	0	3.6	67.8	28.6	0	0	21.4	57.2	21.4	0
TASP – Placement	0	42	31.2	16.6	8.2	2	8	88	4	0
Mass.										
MCAS – Gr 10	0	24	52	24	0	0	4	51	32	13
MCAS retest	0	33	27	20	2	0	3	55	33	8
National admissions & placement										
SAT 1	0	40	32	18	8	2	3	75	7	15
ACT	0	33	30	13	23	0	10	85	0	5
Accuplacer (plcmnt)*	0	17*	74*	0*	9*	0	9	88	3	0

***Elementary algebra test only. This study did not look at the Accuplacer test for college algebra.**

Categories:

Grade level: the point at which the item content is typically taught and learned
Kind of Knowledge: the cognitive ability that is the primary target of the time

SEE "Interpreting cross-state findings", page 4. Also see, "Mathematics Protocol" for a complete description of categories.

Gap Analysis – Cross-State View
Mathematics Content
(in percent of total items)

	Item Content								
	Number	Measure-ment	Geom	Alg & Func.	Calc	Stats	Discrete	Reason	Other
Indiana:									
Core 40 Alg 1	5	0	5	74	0	0	0	16	0
Core 40 Geom	0	26	58	13	0	0	0	3	0
Core 40 Alg 2	6	0	9	74	0	3	3	6	0
Purdue Plcmnt	24	0	0	76	0	0	0	0	0
Kentucky:									
KCCT – Gr 11	12	0	26.8	41.5	0	14.6	4.9	0	0
EKU Plcmnt	28.6	0	0	71.5	0	0	0	0	0
KEMTP	20	0	16.7	63	0	0	0	0	0
WKU Plcmnt	16	0	0	84	0	0	0	0	0
Nevada:									
HS Prof Exam	26	8	10	34	0	22	0	0	0
Texas:									
TAKS - HS	3.6	0	28.5	60.7	0	3.6	0	3.6	0
TASP - Plcmnt	12.5	12.5	14.6	52	0	2	0	6.3	0
Mass:									
MCAS – HS	5.7	13	15	37.7	0	26	1.9	0	0
MCAS -retest	27	12	15	27	0	17	0	3	0
National:									
SAT 1	25	5	18	37	0	10	3	2	0
ACT	17	3	20	43	0	5	3	2	7
Accuplacer*	20	0	0	80	0	0	0	0	0

***Elementary algebra test only. This study did not look at the Accuplacer test for college algebra.**

SEE "Interpreting cross-state findings", page 4. Also see "Mathematics Protocol" for a complete description of categories.

Recommendations & issues to consider

Unlocking the substance of the tests students take as they make the transition from high school to college is a crucial first step toward a coherent K-16 system because tests contain the knowledge and skills students are required to demonstrate. Letter grades are useful but can vary from district to district, school to school, and even classroom to classroom. Because of this, letter grades can conceal educational inequities: we know from research, for example, that students in high-poverty schools who earn A's produce the same test scores as students earning C's and D's in more affluent schools.[3] Course requirements and standards share the same ambiguities as student GPAs. They provide some useful information, for sure, but the specific knowledge and skills students are acquiring isn't actually known until students are tested.

Based on the findings of the test gap analysis, we make the following recommendations to the ADP states and raise some key issues that states should consider:

Short term strategies, or how states can begin to use their test data now.

Some of the ADP states have high school tests that already come close to the knowledge and skills that higher education looks for, especially in reading comprehension and writing. Indiana, Massachusetts and Texas, for example, were noteworthy for high level content in reading. Two of these states, Massachusetts and Texas, have already held discussions to explore using state high school test scores for college placement. But even in states where the gap may be wider, the high school tests provide some information that could be useful to higher education. For example, all the states have an on-demand writing assessment in high school which clearly reveals more about students' writing ability than either the ACT or SAT 1, neither of which assesses writing at all.

Using high school assessment data. There is no requirement for *how* higher education uses the data, only that they use it in some capacity as states move toward aligning high school and college expectations. For most of the 1990s, West Virginia administered the Stanford 9 to all of the state's high school students. No one, not even the test publisher, would argue that this assessment is a measure of college readiness. Nonetheless, the state monitored student progress in state colleges and universities based on Stanford 9 scores and found a direct relationship between high scores and success in college. The ADP states should examine their own data to discover if similar relationships can be found with their tests so that higher education can reinforce existing K-12 standards even while their states move toward a better alignment.

Oregon offers a slightly different image for how this might work. Several years ago, state higher education faculty agreed on five major skills or domains of mathematics that students needed in order to be ready for success in college. The state high school test assesses only three of these five domains, so the K-16 alignment is not airtight. However, higher education uses students' results on the high school test for the three domains the sectors have in

[3] U.S. Department of Education, Office of Educational Research and Improvement, "What Do Student Grades Mean? Differences Across Schools," Education Research Report, January 1994.

common. Moreover, students who perform well enough on the high school exam receive a letter of congratulations, informing them they have already met three of the five math requirements for entering postsecondary.

Some high school tests may already serve as a decent measure of readiness to begin college-level work and may only require a higher score for placement than the state requires for high school graduation. The City University of New York (CUNY) waives the writing placement tests for students who pass the Regents' exam with a specified score. As mentioned, Texas is also moving in this direction and Massachusetts is exploring the option. Higher education can send a strong message in support of K-12 standards merely by asking to see the test scores. Students could also be encouraged to volunteer their scores on high school tests to supplement their college application.

Long-term alignment policies.

Even if current tests provide some meaningful information for colleges, states will still want to align their testing system across sectors. Some students will find a way to rise to the top of any measure. But most students will perform to the levels they are taught. If curriculum and teaching are targeted toward an assessment that falls a long way short of college readiness, these students will arrive at postsecondary institutions lacking the knowledge and skills they need to succeed. It's in students' interests, and thus in the state's interest, to make sure that the assessments used for high school accountability share enough coherence with the core requirements of college admissions and placement so students aren't blindsided when they enter campus for the first time. States will also accrue a financial benefit by reducing the considerable price tag associated with postsecondary remediation.

Remember that the gaps work in both directions. To be sure, the college tests tended to address higher level content than the high school exams did. But most of the high schools exams addressed a wider range of content than either the college admissions or placement tests. This was most apparent in mathematics. Both high school and college exams emphasized algebra and functions, but the high school tests further included statistics and geometry concepts—mathematics that were mostly absent from the college tests. Gaps in writing were also apparent in both sectors, and college placement tests were often less sophisticated than the high school exams. It would also seem that the designers of college placement tests have much to learn from several high school assessments about how to structure items that measure students' conceptual understanding and require students to be more analytical as demanded by open-response items.

An aligned system is meaningless if a weak high school assessment is aligned to a weak college test. States must make sure that their alignment strategies are attached to content and skills that show college readiness. And while pursuing the goal to help students meet higher education's requirements, states must also be careful not to leave some essential content behind. Even if college and universities aren't explicitly looking for knowledge of statistics and geometry, for example, states are likely still interested in knowing that schools are teaching these domains.

Other ways to fill the gap. Some of the answers to filling in the gap may already be present in state standards documents. The standards-mapping provided by the Diploma project will add another valuable piece to the gap-closing puzzle.

States should also review their course requirements for high school graduation. Early testing policies for high school exit exams can explain some of the gap between high school and college tests; after all, it's difficult for states to assess students on content that they have not yet been taught. But states still need to know that students are learning this content before they arrive at college. Aligning course requirements is another way to fill in the gap.

This test gap analysis has revealed the knowledge and skills that need to be bridged between the educational sectors in each of the ADP partner states. Other research—the standards-mapping and the workplace studies—give states even more material for filling the gap, and the legal research helps states imagine what is possible. But ultimately, the success of the American Diploma Project will be determined by the strength of the K-16 conversations that occur in the ADP states.

Report prepared by
Patte Barth, Senior Associate
The Education Trust, Inc.
July 2002